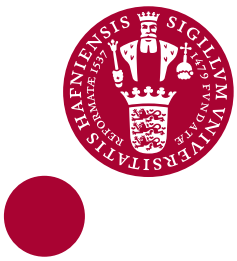

Compact Algorithms for Predicting of Atmospheric Visibility Using PM2.5, Relative Humidity and NO2

Yi, Hui; Zhang, Jingjing; Xiao, Hang; Tong, Lei; Cai, Qiuliang; Lin, Jiamei; Yu, Weijia; Johnson, Matthew S.

Published in:
Aerosol and Air Quality Research

DOI:
10.4209/aaqr.2019.06.0286

Publication date:
2020

Document version
Publisher's PDF, also known as Version of record

Special Session on Optical Properties of Atmospheric Aerosol — Observation, Measurement Techniques and Model Analysis for Improving the Accuracy of Aerosol Light Absorption Determinations in Polluted Sites

Aerosol and Air Quality Research, 20: 679–687, 2020

ISSN: 1680-8584 print / 2071-1409 online
doi: 10.4209/aaqr.2019.06.0286

Compact Algorithms for Predicting the Atmospheric Visibility Using $PM_{2.5}$, Relative Humidity and NO_2

Hui Yi[1,2,3#], Jingjing Zhang[1,2,3#], Hang Xiao[1,3*], Lei Tong[1,3], Qiuliang Cai[1,2,3], Jiamei Lin[1,2,3], Weijia Yu[4], Matthew S. Johnson[4]

[1] *Center for Excellence in Regional Atmospheric Environment & Key Lab of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen 361021, China*
[2] *University of Chinese Academy of Sciences, Beijing 100049, China*
[3] *Ningbo Urban Environment Observation and Research Station, Chinese Academy of Sciences, Ningbo 315800, China*
[4] *Department of Chemistry, University of Copenhagen, Copenhagen Ø 2100, Denmark*

ABSTRACT

Visibility is a key parameter of the atmospheric environment that has attracted increasing public attention. Despite its importance, very few descriptions of methods for predicting visibility using widely available information in the literature exist. In this paper, we derive and evaluate two compact algorithms (Models I and II) for measuring and predicting visibility using records of $PM_{2.5}$, relative humidity (RH) and NO_2 from 16 cities around the world. Models I and II are simplified algorithms derived from Pitchford's algorithm. Our analysis shows that Model I is more consistent with the observations and can accurately predict changes in visibility. In a separate part of the study, the two algorithms are trained using data sets from individual cities. Better results are obtained when the models are trained with the data from London, Sydney and the Chinese mainland cities. Model II displays broader applicability when it is simulated using a single city's data set. This study indicates that atmospheric visibility can be well quantified based on measurements of $PM_{2.5}$, RH and NO_2 concentrations.

Keywords: Atmospheric visibility; Light extinction coefficient; Algorithm; $PM_{2.5}$; Relative humidity; NO_2.

INTRODUCTION

Atmospheric visibility is closely related to daily life. Low visibility can lead to traffic accidents, flight delays and visual impairment, which has attracted more and more public attention. Visibility is easily measured using laser radar, the photograph processing method and aerosol sampling method (Luo *et al.*, 2005), but theoretical algorithms for quantifying and predicting visibility have only rarely received attention. Increased awareness about the negative impacts of visibility on human daily activity motivated the international community to develop new tools for visibility prediction.

Using Koschmieder's formula (Larson and Cass, 1989; Che *et al.*, 2006), $\sigma_{ext} = -\ln 0.02/V$, horizontal visibility (V, km) can be calculated by the atmospheric extinction coefficient (σ_{ext}, km^{-1}). Theoretical descriptions of σ_{ext} have been derived by many people. A simple algorithm for estimating σ_{ext} from measured species concentrations was developed by Malm (1994). However, the accuracy and precision of the algorithm was very low. Later, based on a continuous monitoring campaign of 160 sites through the Interagency Monitoring of Protected Visual Environments Particle Monitoring Network (IMPROVE), a new algorithm for estimating light extinction was developed by Pitchford *et al.* (2007). The algorithm includes 15 variables (such as small sulfate, large sulfate, small nitrate, large nitrate, small organic mass, large organic mass, soil dust and sea salt). The Pitchford model is more consistent with the atmospheric aerosol literature and reduces bias at extremes of high and low light extinction. The improved performance of this model demonstrates that the prediction of visibility involves using data obtained by monitoring meteorological conditions and airborne pollutants. But there are plenty of variables in the model which contribute to lower applicability.

In this paper, we present two compact algorithms (Model I and Model II) to predict visibility based on data for $PM_{2.5}$, relative humidity and NO_2 concentration in 16 cities around the world. As $PM_{2.5}$, relative humidity and NO_2 are easy to measure and predict, visibility can also be predicted by measurement of $PM_{2.5}$, relative humidity and NO_2 using the constructed algorithm. It is of great significance for reducing the cost of measuring visibility in the region where only

[#] The two authors contributed equally to this paper.

[*] Corresponding author.
Tel.: 86-574-8678-4813; Fax: 86-574-8678-4813
E-mail address: hxiao@iue.ac.cn

conventional pollutants and meteorological parameters are monitored.

MATERIALS AND METHODS

Model Assumption

Koschmieder (Che *et al.*, 2006) established an algebraic relationship between visibility and the light extinction coefficient (σ_{ext}), which can be described as $\sigma_{ext} = -\ln 0.02/V = \sigma_{sp} + \sigma_{ap} + \sigma_{sg} + \sigma_{ag}$. The atmospheric light extinction coefficient is the sum of particle scattering coefficient (σ_{sp}), particle absorption coefficient (σ_{ap}), Rayleigh scattering coefficient (σ_{sg}) and gas absorption coefficient (σ_{ag}). Light scattering and absorption by particles are the main reason for visibility degradation; it is directly affected by meteorological factors and airborne contaminants such as nitrate, sulfate, EC, OC, ammonium salt and secondary organic aerosol (Li *et al.*, 2018; Yu *et al.*, 2019). Particles' contribution to the atmospheric extinction coefficient exceeds 95% (Cao *et al.*, 2012). Rayleigh scattering refers to the scattering of light from air molecules, and it depends on the density of atmosphere. Generally, it was assumed to be a constant value of 0.01 km^{-1} at sea level (Watson *et al.*, 2002). Gas absorption is mainly contributed by NO_2 which has small effect on the light extinction coefficient.

Pitchford *et al.* (2007) have developed a precise algorithm (Eq. (1)) for calculating the light extinction caused by different processes:

$$\begin{aligned}-\ln 0.02/V = {} & 2.2f_S(\mathrm{RH})[\text{Small Sulfate}] + \\ & 4.8f_L(\mathrm{RH})[\text{Large Sulfate}] + 2.4f_S(\mathrm{RH})[\text{Small Nitrate}] + \\ & 5.1f_L(\mathrm{RH})[\text{large Nitrate}] + 2.8[\text{Small Organic Mass}] + \\ & 6.1[\text{Large Organic Mass}] + 10[\text{Elemental Carban}] + \\ & [\text{Soil dust}] + 0.33\rho(NO_2) + 1.7f_{SS}(\mathrm{RH})[\text{Sea Salt}] + \\ & 0.6[\text{Coarse Mass}] + \text{Rayleigh Scattering (Site Specific)}\end{aligned} \quad (1)$$

where RH (%) is relative humidity of ambient air; f_S(RH), f_L(RH) and f_{SS}(RH) represent hygroscopic growth functions of small, large aerosol components and sea salt, respectively (all are unitless); $\rho(NO_2)$ represents the atmospheric mass concentration of NO_2 (μg m^{-3}). The other parameters, such as organic mass, sulfate, nitrate, elemental carbon, soil dust and so on, are specific chemical species in the aerosol, which are either obtained by direct measurements or estimation from other experimental parameters based on relevant references (Taylor and McLenna, 1985; Cao *et al.*, 2005, 2012). The large and small fractions such as sulfate, nitrate and OM indicate different formation process through dry and aqueous mechanisms (John *et al.*, 1990), which are calculated according to Cao *et al.* (2012). Component concentrations shown in brackets are in μg m^{-3}. Despite the complexity of the parameters and manifestations of this equation, it can be summed up in two basic assumptions: 1) The total extinction effect of the atmosphere can be approximated as a combination of the extinction effects of the chemical components in dry air and the hygroscopic growth; 2) the extinction effect of each component follows Beer-Lambert law.

However, the use of the algorithms requires many parameters, which hamper its applicability. Therefore, the main purpose of this paper is to simplify Pitchford's algorithm. In Pitchford's algorithm, the light extinction coefficient is the sum of particle scattering/absorption coefficient, Rayleigh scattering coefficient and gas absorption coefficient. Sulfate, nitrate, organic mass, elemental carbon, sea salt and coarse mass were the main hygroscopic aerosol components which contributed to the particle scattering/absorption coefficient. The particle scattering/absorption coefficients of above species were calculated separately. In the simplified algorithms, the hygroscopic aerosol components are a portion of $PM_{2.5}$ which have been replaced by $PM_{2.5}$. The water growth function [f(RH)] was replaced by aerosol humidification factor $(1 - RH/100)^b$. $PM_{2.5}$ is directly multiplied by $(1 - RH/100)^b$ in Model I. Regression analysis between $PM_{2.5}$ and visibility has indicated that exponential equations can describe their algebraic relationship (Chan *et al.*, 1999; Zhang *et al.*, 2019). So $PM_{2.5}$ was exponential conversed by $\exp(PM_{2.5}/e_2)$ in Model II. Rayleigh scattering is a site-specific parameter that contributes less than 2% to the value of the extinction coefficient (Cao *et al.*, 2012). Therefore, this term was omitted. The last term of $\rho(NO_2)$ represents the gas absorption effect which is retained in the simplified algorithms. Based on these assumptions, two models were constructed and presented in Eqs. (2) and (3):

Model I:

$$\sigma_{ext} = \frac{-\ln(0.02)}{V} = a_1 \times \mathrm{PM}_{2.5} \times \left(1 - \frac{RH}{100}\right)^{b_1} + c_1 \times \rho(\mathrm{NO}_2) + d_1 \quad (2)$$

Model II:

$$\sigma_{ext} = \frac{-\ln(0.02)}{V} = a_2 \times \exp\left(\frac{\mathrm{PM}_{2.5}}{e_2}\right) \times \left(1 - \frac{RH}{100}\right)^{b_2} + c_2 \times \rho(\mathrm{NO}_2) + d_2 \quad (3)$$

where $PM_{2.5}$ represents the mass concentration of fine particles (in μg m^{-3}); $\rho(NO_2)$ represents the atmospheric mass concentration of NO_2 (in μg m^{-3}); a_1 and a_2 represent the coefficient of the $PM_{2.5}$; b_1 and b_2 describe the coefficients of humidification factor; c_1 and c_2 represent the coefficient of NO_2; e_2 is the correction factor of $PM_{2.5}$; d_1 and d_2 are the error terms. The unknown coefficients of Model I (a_1, b_1, c_1 and d_1) and Model II (a_2, b_2, c_2, d_2 and e_2) are determined using iterative regression to fit the data from the 16 cities.

Data Collection and Data Processing

In order to derive a robust compact model, data for visibility, $PM_{2.5}$, relative humidity and NO_2 concentrations from 16 large cities around the world (Beijing, Guangzhou, Hangzhou, Ningbo, Xiamen, Shijiazhuang, Chongqing, Shanghai, Pinzhen, Xinbei, London, Sacramento, Toronto, New York, Coyhaique and Sydney) were downloaded from the Center of National Ministry of Environmental Protection of China (MEPC) (http://datacenter.mep.gov.cn/) and the OpenAQ air quality database (https://openaq.org). These data were processed by standardized procedures (Doyle *et al.*, 2002;

Che *et al.*, 2008; Ratto *et al.*, 2012): (1) All variable units are converted into standard units; (2) the outlying data points are removed before regression; (3) the visibility values should be greater than 0.3 km, and less than the 99th percentile of the data; (4) data obtained during abnormal weather episodes such as precipitation, mist and dust storms were excluded. After the screening process, a total of 10,107 points have been retained for regression.

Model Regression

The original data was divided into two parts; the main part of the data set was used for model simulation and the remainder was used to validate the model. First, the unknown coefficients of Model I and Model II were obtained through iterative regression with all of the 16 cities' data. The regression analyses were performed using the 1stOpt1.5 software package (Liu *et al.*, 2018; Marques *et al.*, 2019). 1stOpt1.5 is an automatically running optimized regression software which is developed by 7D Soft High Technology Inc. (China). When an equation with undetermined coefficients is inputted to the program, the software will automatically perform iterative regression until a converged solution is obtained thereby determining the best coefficient values. After determining the model parameters, sensitivity analysis was conducted using Monte Carlo random simulation method for the atmospheric visibility estimation.

In order to compare the pros and cons of the two algorithms, the values of Akaike information criteria (AIC) and Bayesian information criteria (BIC) were calculated using Origin 2017. In statistics, the Akaike information criterion is an estimator of the relative quality of statistical models for a given set of data. It is based on the likelihood function. Given a collection of models for the data, AIC estimates the quality of each model, relative to each of the other models. Thus, AIC provides a means for model selection. Similar to Akaike information criteria, the Bayesian information criterion is also a criterion for model selection among a set of models. The model with the lowest AIC and BIC is preferred (Akaike *et al.*, 1998; McDonald *et al.*, 2016). The calculation formulas for AIC and BIC are shown in Eqs. (4) and (5) (Vrieze, 2012):

$$k_{AIC} = 2\kappa - 2l_Y(\hat{\tau}) \tag{4}$$

$$k_{BIC} = \kappa lg(N) - 2l_Y(\hat{\tau}) \tag{5}$$

where k_{AIC} and k_{BIC} represent the AIC and BIC values (unitless), respectively; κ represents the number of estimated model parameters (unitless); $\hat{\tau}$ represents the optimized model parameters; $l_Y(\hat{\tau})$ is the log of the likelihood of $\hat{\tau}$ given the data of Y; N represents the number of observations (unitless).

RESULTS AND DISCUSSION

Visibility Distribution

The aerosol contribution to "global dimming" was first reported as a strong decrease in visibility up to the middle 1980s. Since that time, visibility has increased over Europe, consistent with reported European "brightening", but has decreased substantially over south and east Asia, South America, Australia, and Africa, resulting in net global dimming over land (Wang *et al.*, 2009; Vautard *et al.*, 2009).

Statistical summaries of the atmospheric visibility, $PM_{2.5}$, RH and $\rho(NO_2)$ are shown in Table 1. During the study period, low visibility (less than 10 km) was observed in Hangzhou (7.86 km), Guangzhou (8.89 km), Chongqing (7.27 km), Shanghai (7.12 km) and Xinbei (6.16 km) with high $PM_{2.5}$ levels (36.39–58.73 μg m^{-3}). It is well known that $PM_{2.5}$ is the main factor which contributes to visibility degradation. The visibility in Beijing (11.0 km), Shijiazhuang (10.7 km) and Coyhaique (12.2 km) were above 10 km, which exhibited higher $PM_{2.5}$ concentrations compared with those in the low-visibility cites. It indicates that visibility is also affected by other conditions. Visibility varies with the degree of air pollution. Most Chinese cities are facing a downward trend in visibility due to the high $PM_{2.5}$ level (Che *et al.*, 2008; Molnar *et al.*, 2008). Some cities such as Pinzhen, Sacramento, New York and Sydney exhibited relatively good visibility (above 15 km). The mean concentrations of $PM_{2.5}$ (7.33–16.14 μg m^{-3}) and $\rho(NO_2)$ (7.2–23.9 μg m^{-3}) were also very low. In London, due to vehicle emissions including a greater fraction of diesel engines, a relatively high NO_2 concentration of 88.9 μg m^{-3} was observed. However, NO_2 has a very small effect on visibility. On the other hand, the $PM_{2.5}$ concentration of 15.1 μg m^{-3} was very low, and the visibility is also above 15 km.

Model Simulation with Combined Data

The coefficients for the two models were determined using all of the 16 cities' data by Levenberg-Marquardt and Universal Global Optimization methods (convergence criterion: 1.00E-10; maximum iterations number: 1000; repeat number: 30; control iterations number: 20). The running codes for each model were shown in Table 2. Both Model I and Model II achieved convergence after 17 and 40 iterations. The obtained equations are listed in Table 3. The adjusted R^2 of Model I (0.58) is a little higher than Model II (R^2 = 0.57) and the AIC value of 26,909 and BIC value of 26,945 were smaller than the corresponding values in Model II. Preliminary inspection indicated that Model I is the optimal algorithm for visibility prediction.

Sensitivity analysis was conducted using Monte Carlo random simulation method for the atmospheric visibility estimation (Fig. 1). It shows that the combination of most significant contributors, $PM_{2.5}$ and RH, could account for 97–98% of the visibility degradation. In particular, the contributions of $PM_{2.5}$ to the variance of visibility were 72% and 66% from Model I and Model II prediction, respectively. In contrast, the concentration of NO_2 showed a negligible influence on the variance of visibility, with uncertainty contributions of 2% and 3% from Model I and Model II, respectively. The sensitivity analysis indicated that decreasing the $PM_{2.5}$ concentration and RH can significantly improve atmospheric visibility.

Model I and II are four-dimensional functions; it is hard to directly draw the function surface. However, we note that NO_2 has a limited influence on visibility; the main contributors

Table 1. Sites information and the values of atmospheric visibility, $PM_{2.5}$, RH and $\rho(NO_2)$ in the 16 cities.

City[a]	Country/Region	Latitude, Longitude	Visibility (km)		$\rho(NO_2)$ (μg m^{-3})		$PM_{2.5}$ (μg m^{-3})		RH (%)	
			Mean	Range	Mean	Range	Mean	Range	Mean	Range
Beijing	China	40.0801N, 116.5846E	11.0	1.3–30.0	70.04	3.00–374.96	53.8	8.6–89.9	39.0	0.5–110.1
Hangzhou	China	23.3924N, 113.2988E	7.9	0.6–20.0	52.59	7.28–269.51	65.1	29.3–89.8	27.9	5.2–100.1
Guangzhou	China	30.2232N, 120.1325E	8.9	1.5–22.9	36.39	3.41–140.00	76.8	39.1–98.7	37.7	8.4–123.1
Ningbo	China	29.7325N, 121.6511E	12.0	1.2–30.0	45.46	5.29–389.78	71.1	31.8–89.8	29.9	0.5–98.3
Xiamen	China	24.5440N, 118.1277E	12.4	3.4–26.9	26.34	5.50–96.79	76.9	40.0–90.0	32.7	5.3–87.9
Shijiazhuang	China	38.0667N, 114.3515E	10.7	0.4–30.0	90.61	5.00–623.48	54.4	10.7–90.0	56.4	3.4–225.7
Chongqing	China	29.7192N, 106.6417E	7.3	1.4–23.9	55.07	8.07–255.42	72.8	38.8–90.0	46.0	2.5–121.0
Shanghai	China	31.1433N, 121.8052E	7.1	2.2–10.0	58.63	6.29–236.15	78.7	37.1–89.9	42.0	7.5–131.4
Xinbei	Taiwan	25.0333N, 121.5167E	6.2	1.2–9.0	37.13	4.00–85.00	73.9	39.4–89.0	20.2	2.4–54.0
Pinzhen	Taiwan	25.0543N, 121.6325E	16.9	1.0–60.0	16.14	2.00–77.00	69.8	41.7–90.0	16.7	1.4–43.0
London	UK	51.4775N, 0.4614W	15.4	3.2–31.7	15.10	3.27–73.36	74.5	47.0–90.0	88.9	23.0–189.4
Toronto	Canada	43.6167N, 79.3833W	13.5	5.1–15.0	7.64	1.35–33.36	71.9	39.0–89.9	15.9	3.4–51.9
New York	US	40.6386N, 73.7622W	15.0	0.9–16.1	7.33	1.13–28.84	62.5	21.1–89.8	16.9	2.9–59.1
Sacramento	US	38.5833N, 121.0500W	15.1	6.4–16.1	15.69	1.00–66.00	52.4	8.5–89.6	7.2	1.0–36.0
Coyhaique	Chile	45.5942S 72.1061W	12.2	0.4–30.0	67.27	6.79–469.40	78.0	29.2–100.0	23.1	1.9–105.5
Sydney	Australia	33.9461S 151.1772E	17.6	6.4–40.0	8.07	1.90–49.40	65.7	17.6–90.0	23.9	3.8–57.0

[a] Study period. Beijing: 2015.1.1–2016.12.31; Hangzhou: 2015.1.1–2016.12.31; Guangzhou: 2013.1.1–2013.12.31; Ningbo: 2013.6.1–2015.5.31; Xiamen: 2015.1.1–2016.12.31; Shijiazhuang: 2015.1.1–2016.12.31; Chongqing: 2015.1.1–2016.12.31; Shanghai: 2015.1.10–2016.9.9; Coyhaique: 2017.8.14–2017.9.20; Xinbei: 2017.9.12–2017.12.25; London: 2015.1.1–2016.12.31; Pinzhen: 2017.9.12–2018.2.22; Sacramento: 2017.8.19–2018.4.9; Toronto: 2015.1.1–2016.12.31; New York: 2015.1.1–2016.12.31; Sydney: 2015.1.1–2017.11.8.

Table 2. The running codes for model regression by 1stOpt1.5 software.

Name	Code
Model I	Title: Model I; Parameters: a, b, c, d; Variable: V, R, P, N; Function V = –ln0.02/(a × P × ((1 – R × 0.01)^{b}) + c × N + d); Data: 13.59[a] 75.66[b] 1.90[c] 17.10[d] 10.10 85.15 2.10 19.00 12.38 75.00 2.10 11.40 ⋮ ⋮ ⋮ ⋮
Model II	Title: Model II; Parameters: a, b, c, d, e; Variable: V, R, P, N; Function V = –ln0.02/(a × exp(P/e) × ((1 – R × 0.01)^{b}) + c × N + d); Data: 13.59 75.66 1.90 17.10 10.10 85.15 2.10 19.00 12.38 75.00 2.10 11.40 ⋮ ⋮ ⋮ ⋮

[a] The 1st column is visibility. [b] The 2nd column is relative humidity. [c] The 3rd column is $PM_{2.5}$. [d] The 4th column is $\rho(NO_2)$. V: visibility, km; R: relative humidity, %; P: $PM_{2.5}$, μg m^{-3}; N: $\rho(NO_2)$, μg m^{-3}.

Table 3. The functions of the two models in this study and that used in Pitchford *et al.* (2007).

Name	Algorithm	R^2	k_{AIC}	k_{BIC}
Model I	$\sigma_{ext} = -\ln 0.02/V$ $= 0.00143PM_{2.5}(1 - 0.01RH)^{-1.10731} - 0.00073\rho(NO_2)$ $+ 0.21376$	0.58	26,909	26,945
Model II	$\sigma_{ext} = -\ln 0.02/V$ $= 0.06561\exp(PM_{2.5}/50.19809) \times (1 - 0.01RH)^{-0.59352}$ $- 0.00075\rho(NO_2) + 0.11489$	0.57	27,070	27,114
Pitchford	Eq. (1)	0.86[a]	-	-

[a] The R^2 correlation coefficient of 0.86 is obtained from reference using IMPROVE particle speciation data.

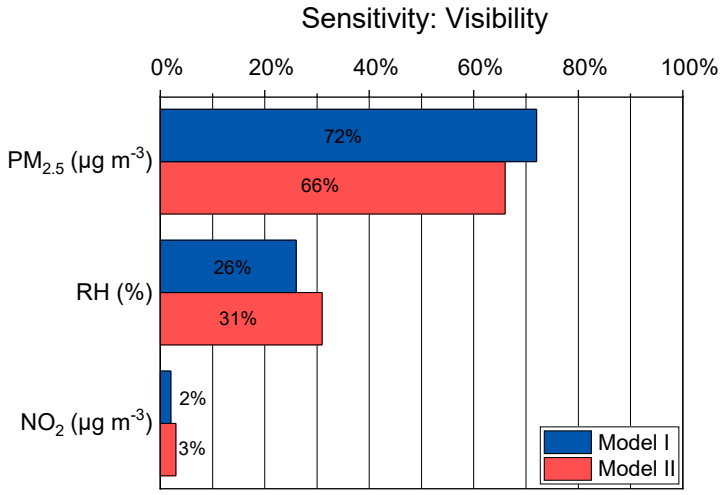

Fig. 1. Sensitivity analyses of $PM_{2.5}$, RH and NO_2 on atmospheric visibility predicted by Model I and Model II.

for visibility degradation are $PM_{2.5}$ components (Tao *et al.*, 2007). For the purpose of discussion, $\rho(NO_2)$ was omitted and the models become three-dimensional. The three-dimensional surfaces of Model I and II are shown in Fig. 2. The surfaces are extended upwards with visibility increasing with decreasing $PM_{2.5}$ and humidity. The Model I surface is in the middle of the scattered data points which indicate a satisfactory simulation result. However, it should be noted that the uncertainty

contribution of $\rho(NO_2)$ from Model II (3%) is slightly higher than that from Model II (2%). Without considering the contribution of $\rho(NO_2)$, the surface of Model II deviates to a slightly bigger extent from the dense area compared with the surface of Model I. Both Model I and II have good predictive results for visibilities less than 20 km. Large deviations appear when visibility is higher than 20 km. It can be seen that the models derived from the combined data of 16 cities have certain generality, especially given the different atmospheric environments represented within the data set. An evaluation for the predictive models obtained with combined data by the single-city data was performed, and the result (Table S1) showed that the adjusted R^2 of the two models in Chinese mainland cities (Beijing, Guangzhou, Hangzhou, Ningbo, Xiamen, Shijiazhuang, Chongqing, Shanghai) and London are higher than those obtained in other cities. The prediction results in most cites are satisfied, which suggested the adaptability of these two models.

In comparison, the R^2 in the algorithm of Pitchford (Table 3), with the value of 0.86, was obtained from the IMPROVE particle speciation data, which is higher than the results obtained in this study. The Pitchford algorithm is theoretical

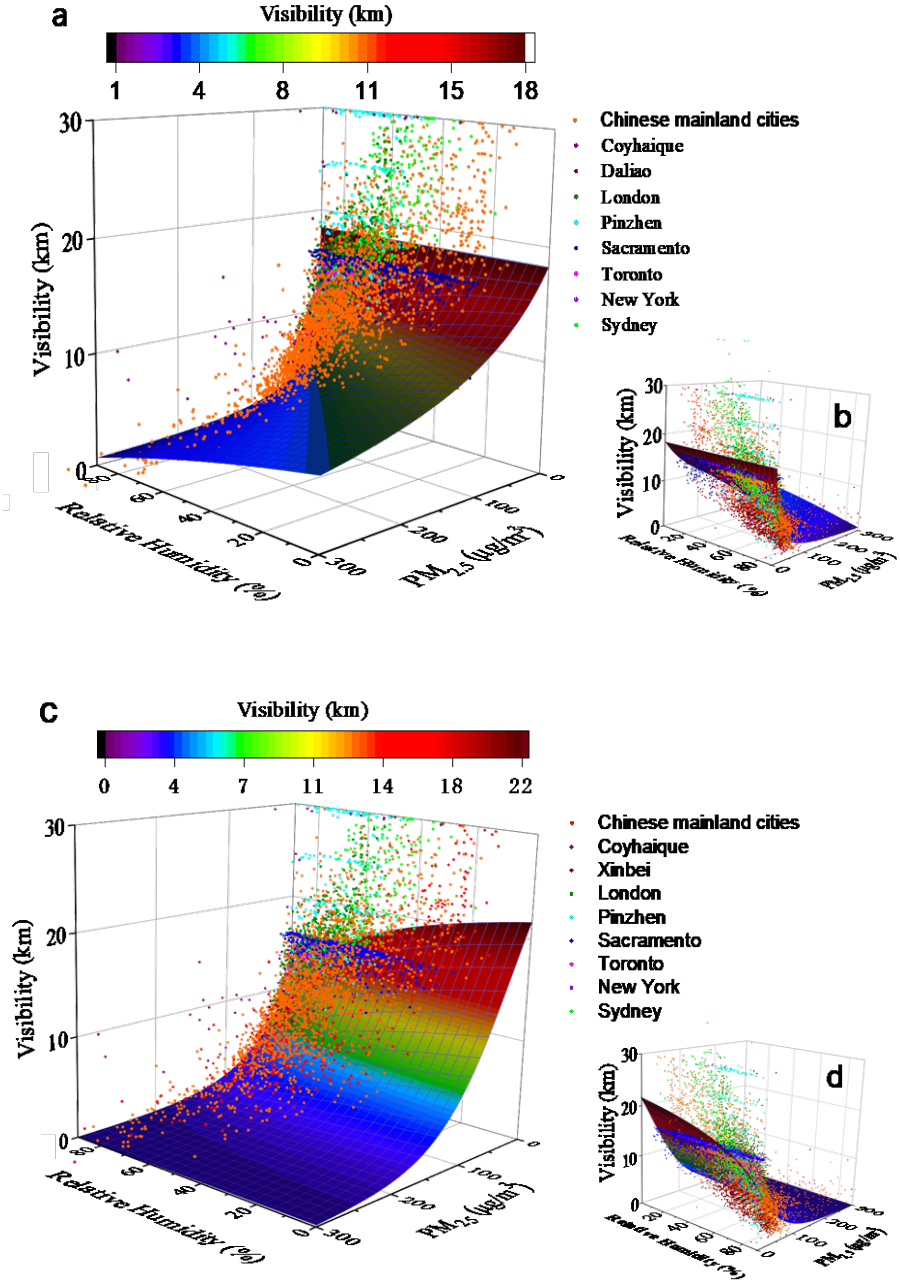

Fig. 2. Three-dimensional surface diagrams of Model I and II based on fits to the data from the 16 cities: (a) Model I, (b) counterclockwise rotation of the Model I surface by 180°, (c) Model II, (d) counterclockwise rotation of the Model II surface by 180°. Scatter points in different colors represent the observed data from different cities.

prediction of visibility with 15 entangled variables, which may over-describe the system (Kelly *et al.*, 2013). While the algorithms derived in this study only have 3 variables ($PM_{2.5}$, RH and $\rho(NO_2)$), the advantage is that there are fewer variables and these variables are easier to obtain. The estimated visibility value is acceptable for many purposes.

Simulation with Single-city Data

The equations of Model I and II were derived with single-city data using 1stOpt software, following the same protocol. The values of the coefficients for the two models for each city are shown in Tables S2 and S3. The adjusted R^2 (Table 4) of the two models in Beijing, Guangzhou, Hangzhou, Ningbo, Xiamen, Shijiazhuang, Chongqing, Shanghai and London are higher than other cities (Model I: 0.62–0.86; Model II: 0.65–0.87; $p < 0.01$). The results are comparable to those obtained by Pitchford *et al.* (2007). Compared with Model I, Model II has higher R^2 in many cities such as Guangzhou, Xiamen, Coyhaique, Xinbei, Pinzhen, Sacramento, Toronto, New York and Sydney. It indicated that Model II has broader applicability when simulated with a single-city data set. Both Model I and Model II have exhibited lower adjusted R^2 (0.19–0.41) in Coyhaique, Xinbei, Pinzhen, Sacramento, Toronto and New York. Using the models to predict the visibility of the above cities is not very accurate. The visibility of these cities might be related to other factors (e.g., unique aerosol composition, and fraction of $PM_{2.5}$ in PM_{10}) in addition to $PM_{2.5}$, relative humidity and NO_2 concentration (Park *et al.*, 2018). Besides, it can be found that the two models derived from the single-city data show higher adjusted R^2, which were better than those obtained with combined data.

Comparisons between model-predicted visibility and observed visibility of these cities are shown in Fig. 3. The results in London and mainland China cities were better than other cities. Xinbei and Pinzhen are located in Taiwan which are near mainland China, but the predicted results were not satisfactory. The prediction results of Sacramento, Coyhaique and Toronto have large deviations due to the small data volume.

CONCLUSIONS

This study proposes two algorithms using $PM_{2.5}$, RH and NO_2 as independent variables for simulating the visibility. According to the simulation based on the combined data of 16 cities, Model I exhibits slightly better applicability. Model II displays broader applicability when it is simulated using a single city's data set. The model predictions are satisfactory for Beijing, Guangzhou, Hangzhou, Ningbo, Xiamen, Shijiazhuang, Chongqing, Shanghai and London (Model I: $R^2 = 0.62–0.86$; Model II: $R^2 = 0.65–0.87$). Lower adjusted values for R^2 (0.19–0.41) are obtained for Coyhaique, Xinbei, Pinzhen, Sacramento, Toronto and New York.

The simulation results confirm that flexible and simpler algorithms can generally produce reliable predictions based on measurements of $PM_{2.5}$, RH and $\rho(NO_2)$. As far as we know, this is the first study to propose simplified algorithms with only 3 variables for visibility prediction. Some limitations in the performance should be noted, partly because the amount of data available for some of the cities was insufficient. Increasing the volume of the data set is necessary in order to improve the models' adaptability and correlatability in future work.

FUNDING

This work was supported by the National Natural Science Foundation of China (No. U1405235), Science and Technology Plan Project of Ningbo City (No. 2015C110001) and Natural Science Foundation of Ningbo City (No. 2015A610247).

Table 4. The correlation coefficients and RSS results of the two models based on city-specific data with observations for each city.

City	Model I		Model II		N
	R^2	RSS	R^2	RSS	
Beijing	0.70	7538.1	0.72	6910.6	557
Hangzhou	0.86	1505.3	0.87	1278.4	514
Guangzhou	0.74	1307.1	0.76	1206.6	342
Ningbo	0.74	4807.9	0.72	5292.8	619
Xiamen	0.66	2158.0	0.68	2059.9	395
Shijiazhuang	0.77	8547.0	0.77	8780.6	692
Chongqing	0.82	1403.3	0.81	1445.5	682
Shanghai	0.74	345.9	0.73	354.5	385
Xinbei	0.38	3156.0	0.39	3112.9	1198
Pinzhen	0.28	28,149.2	0.31	26,723.8	460
London	0.62	3893.3	0.65	3573.9	612
Toronto	0.36	1696.2	0.39	1629.7	700
New York	0.35	2165.4	0.41	1950.6	674
Sacramento	0.29	2705.9	0.31	2611.8	1350
Coyhaique	0.19	4545.6	0.38	2977.2	370
Sydney	0.46	13,473.6	0.62	9483.5	557

RSS: residual sum of squares; N: number of data set samples.

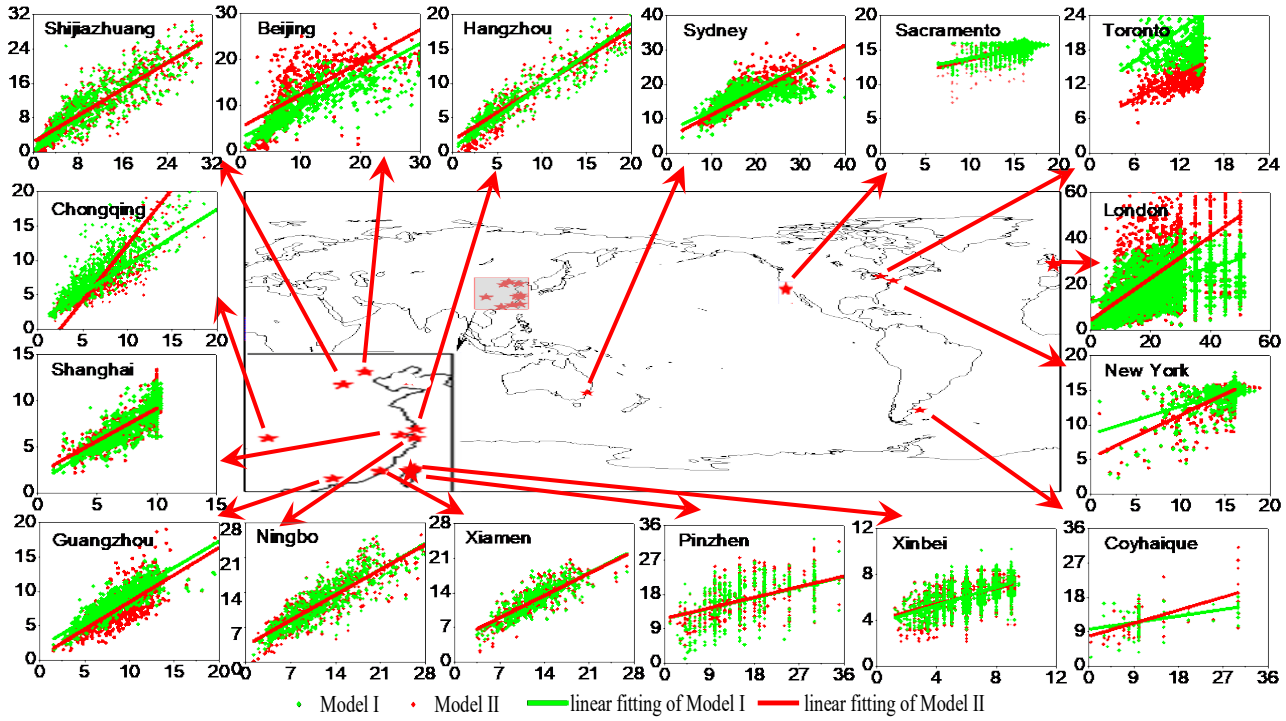

Fig. 3. Comparison between predicted visibility and observed visibility for single-city data simulations.

SUPPLEMENTARY MATERIAL

Supplementary data associated with this article can be found in the online version at http://www.aaqr.org.

REFERENCES

Akaike, H. (1998). *Information Theory and an Extension of the Maximum Likelihood Principle*. Springer Press, New York.

Cao, J.J., Wu, F., Chow, J.C., Lee, S.C., Li, Y., Chen, S.W., An, Z.S., Fung, K.K., Watson, J.G., Zhu, C.S. and Liu, S.X. (2005). Characterization and source apportionment of atmospheric organic and elemental carbon during fall and winter of 2003 in Xi'an, China. *Atmos. Chem. Phys.* 5: 3127–3137.

Cao, J.J., Wang, Q.Y., Chow, J.C., Watson, J.G., Tie, X.X., Shen, Z.X., Wang, P. and An, Z.S. (2012). Impacts of aerosol compositions on visibility impairment in Xi'an, China. *Atmos. Environ.* 59: 559–566.

Chan, Y.C., Simpson, R.W., Mctainsh, G.H., Vowles, P.D., Cohen, D.D. and Bailey, G.M. (1999). Source apportionment of visibility degradation problems in Brisbane (Australia) using the multiple linear regression techniques. *Atmos. Environ.* 33: 3237–3250.

Che, H., Zhang, X., Li, Y. and Zhou, Z. (2006). Relationship between horizontal extinction coefficient and PM_{10} concentration in Xi'an, China. *China Particulolgy* 4: 327–329.

Che, H., Zhang, X., Li, Y., Zhou, Z., Qu, J.J. and Hao, X. (2008). Haze trends over the capital cities of 31 provinces in China, 1981–2005. *Theor. Appl. Climatol.* 97: 235–242.

Doyle, M. and Dorling, S.R. (2002). Visibility trends in the UK 1950–1997. *Atmos. Environ.* 36: 3161–3172.

John, W., Wall, S.M., Ondo, J.L. and Winklmayr, W. (1990). Modes in the size distributions of atmospheric inorganic aerosol. *Atmos. Environ.* 24: 2349–2359.

Kelly, R.A., Jakeman, A.J., Barreteau, O., Borsuk, M.E., ElSawah, S., Hamilton, S.H., Henriksen, H.J., Kuikka, S., Maier, H.R., Rizzoli, A.E., Delden, H.V. and Voinov, A.A. (2013). Selecting among five common modelling approaches for integrated environmental assessment and management. *Environ. Modell. Software* 47: 159–181.

Larson, S.M., Cass, G.R. (1989). Characteristics of summer midday low-visibility events in the Los Angeles area. *Environ. Sci. Technol.* 23: 281–289.

Li, Y.C., Shu, M., Ho, S.S.H., Yu, J.Z., Yuan, Z.B., Liu, Z.F., Wang, X.X. and Zhao, X.Q. (2018). Effects of chemical composition of $PM_{2.5}$ on visibility in a semi-rural city of Sichuan Basin. *Aerosol Air Qual. Res.* 18: 957–968.

Liu, H., Wu, B., Liu, S., Shao, P., Liu, X., Zhu, C., Wang, Y., Wu, Y., Xue, Y., Gao, J., Hao, Y. and Tian, H. (2018). A regional high-resolution emission inventory of primary air pollutants in 2012 for Beijing and the surrounding five provinces of North China. *Atmos. Environ.* 181: 20–33.

Luo, C.H., Wen, C.Y., Yuan, C.S., Liaw, J.J., Lo, C.C. and Chiu, S.H. (2005). Investigation of urban atmospheric visibility by high-frequency extraction: Model development and field test. *Atmos. Environ.* 39: 2545–2552.

Malm, W.C. (1994). Spatial and seasonal trends in particle concentration and optical extinction in the United State. *J. Geophys. Res.* 99: 1347–1370.

Marques, G.M., Lika, K., Augustine, S., Pecquerie, L. and

Kooijman, S.A.L.M. (2019). Fitting multiple models to multiple data sets. *J. Sea Res.* 143: 48–56.

McDonald, K.S. (2016). An ecological risk assessment for managing and predicting trophic shifts in estuarine ecosystems using a Bayesian network. *Environ. Modell. Software* 85: 202–216.

Molnar, A., Meszaros, E., Imre, K. and Rull, A. (2008). Trends in visibility over Hungary between 1996 and 2002. *Atmos. Environ*. 42: 2621–2629.

Park, S.M., Song, I.H., Park, J.S., Oh, J., Moon, K.J., Shin, H.J., Ahn, J.Y., Lee, M.D., Kim, J. and Lee, G. (2018). Variation of $PM_{2.5}$ chemical compositions and their contributions to light extinction in Seoul. *Aerosol Air Qual. Res.* 18: 2220–2229.

Pitchford, M., Malm, W., Schichtel, B., Kumar, N., Lowenthal, D. and Hand, J. (2007). Revised algorithm for estimating light extinction from IMPROVE particle speciation data. *J. Air Waste Manage. Assoc.* 57: 1326–1336.

Ratto, M., Castelletti, A. and Pagano, A. (2012). Emulation techniques for the reduction and sensitivity analysis of complex environmental models. *Environ. Modell. Software* 34: 1–4.

Taylor, S.R. and McLenna, S.M. (1985). *The Continental Crust: Its Composition and Evolution*, Blackwell Press. Oxford.

Vautard, R., Yiou, P. and Van Oldenborgh, G.J. (2009). Decline of fog, mist and haze in Europe over the past 30 years. *Nat. Geosci*. 2: 115–119.

Vrieze, S.I. (2012). Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychol. Methods* 17: 228–243.

Wang, K., Dickinson, R.E. and Liang, S. (2009). Clear sky visibility has decreased over land globally from 1973 to 2007. *Science* 323: 1468–1470.

Watson, J.G. (2002). Visibility: Science and regulation. *J. Air Waste Manage. Assoc.* 52: 628–713.

Yu, X.N., Shen, L., Xiao, S.H., Ma, J., Lu, R., Zhu, B., Hu, J.L., Chen, K. and Zhu, J. (2019). Chemical and optical properties of atmospheric aerosols during the polluted periods in a megacity in the Yangtze River Delta, China. *Aerosol Air Qual. Res.* 19: 103–117.

Zhang, J.J., Tong, L., Peng, C.H., Zhang, H.L., Huang, Z.W., He, J. and Xiao, H. (2019). Temporal variability of visibility and its parameterizations in Ningbo, China. *J. Environ. Sci.* 77: 372–382.

Received for review, June 4, 2019
Revised, October 17, 2019
Accepted, February 5, 2020